

OPIC
OFFICE DE LA PROPRIÉTÉ
INTELLECTUELLE DU CANADA



CIPO
CANADIAN INTELLECTUAL
PROPERTY OFFICE

OTTAWA HULL KIA OCJ

(11) (C) 2,045,959

(22) 1991/06/28

(43) 1992/01/03

(45) 1996/04/02

BREVETS

MARQUES
DE COMMERCE

DROITS
D'AUTEUR

DESSINS
INDUSTRIELS

TOPOGRAPHIES
DE CIRCUITS
INTÉGRÉS

PATENTS

TRADE-MARKS

COPYRIGHT

INDUSTRIAL
DESIGN

INTEGRATED
CIRCUIT
TOPOGRAPHY

(51) Int.Cl. ⁶ G10L 5/06

(19) (CA) CANADIAN PATENT (12)

(54) Speech Recognition Apparatus

(72) Hayashi, Haruyuki , Japan

(73) NEC Corporation , Japan

(30) (JP) Japan 2-174801 1990/07/02

(57) 3 Claims



02 AVR. 1996

2045959

ABSTRACT OF THE DISCLOSURE

A speech recognition apparatus for a speech recognition answering system which uses telephone channels and having a function of detecting a PB (PUSH BUTTON) signal. A speech
5 recognition unit recognizes a speech signal from an input signal, while a PB detection unit detects a PB signal from an input signal. A control unit automatically determines whether an input signal is a speech signal or whether it is a PB signal on the basis of the outputs of the speech recognition unit and PB detection
10 unit.

B

SPEECH RECOGNITION APPARATUS

BACKGROUND OF THE INVENTION

The present invention relates to a speech recognition apparatus for use in a speech recognition answering system and, more particularly, to a speech recognition apparatus having a
5 PB (PUSH BUTTON) signal detecting or receiving function.

A conventional speech recognition apparatus of the type described has a speech recognition unit (SRU) and a PB signal recognition unit or PB receiver (PBR), but it cannot determine whether an input signal from a telephone channel is a PB signal
10 or whether it is a speech. It has been customary, therefore, to assign an independent telephone channel to each of PB input and speech input. A business application or similar software using the apparatus monitors the telephone channels to determine which of the telephone channels has received a call and
15 commands the apparatus to use only one of the recognition units SRU and PBR associated with the telephone channel of interest.

Moreover, the conventional speech recognition apparatus with a PB receiving function forces the business application to execute processing matching the independent telephone channels.
20 On the other hand, the user has to select either one of two



2045959

different telephone numbers assigned to speech input and P input, resulting in limited serviceability. In a system wherein the apparatus is expected to call the user, the user has to register desired one of the speech input and PB input at the
5 system beforehand. Further, since the channels and the kinds of input signals are fixedly held in one-to-one correspondence, an idle channel cannot be efficiently assigned. Specifically, when calls concentrate on either one of the speech input and PB input channels, the user cannot take full advantage of the service.

10

SUMMARY OF THE INVENTION

It is therefore an object of the present invention to provide a speech recognition apparatus capable of automatically determining whether an input signal is a speech signal or whether
15 it is a PB signal.

It is another object of the present invention to provide a generally improved speech recognition apparatus.

A speech recognition apparatus of the present invention comprises a speech recognition unit for recognizing a speech
20 from an input signal and outputting the result of recognition, a PB signal detection unit for detecting a PB signal from the input signal and outputting the result of detection, and a control unit for controlling the speech recognition unit and PB signal detection unit to automatically determine whether the input signal is a
25 speech signal or whether it is a PB signal on the basis of the

result of recognition and the result of detection which the speech recognition unit and PB signal detection unit output when used at the same time.

5

BRIEF DESCRIPTION OF THE DRAWINGS

The above and other objects, features and advantages of the present invention will become more apparent from the following detailed description taken with the accompanying drawings in which:

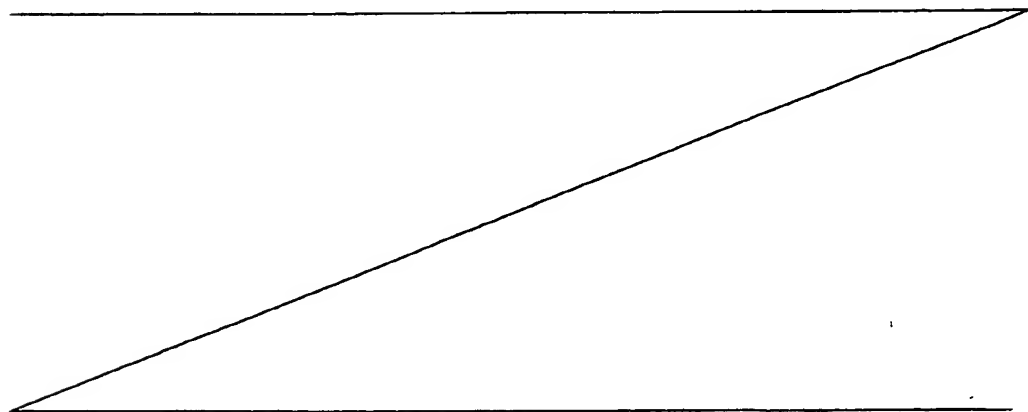
10

FIG. 1 is a block diagram schematically showing a speech recognition system implemented with a speech recognition apparatus embodying the present invention;

FIG. 2 is a flowchart demonstrating a specific operation of the system shown in FIG. 1; and

15

FIGS. 3, 4 and 5A through 5G are flowcharts showing specific procedures which the embodiment executes for automatic speech/PB discrimination.



2045959

DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring to FIG. 1 of the drawings, a speech recognition system implemented with a speed recognition apparatus embodying the present invention is shown. As shown, a subscriber's telephone 6 is connected to a channel control 1 via first and second exchanges 7 and 8. The channel control 1 sends an input signal to a speech recognition apparatus 10 under the control of a business application 2. The speech recognition apparatus 10 has a speech recognition unit (SRU) 3, a PB signal or dial tone recognition unit (PBR), and a control unit 5.

The operation of the system will be described with reference also made to FIG. 2. When a person using the system originates a call on the telephone 6, the call is sent to the channel control 1 via the exchanges 7 and 8 (step 201). In response, the channel control 1 informs the business application 2 of the arrival of the call. On receiving a call termination command from the business application 2, the channel control 1 connects the channel to the speech recognition apparatus 10 (step 202) and then informs the business application 2 of the end of call termination. In response, the business application 2 notifies the control unit 5 of the apparatus 10 of the number of words of an input signal to be recognized (step 203). The number of kinds of input signals to be recognized at the same time are equal to the number of kinds of PB dials of the telephone, and

B

most of them are numerals. Hence, in the illustrative embodiment, let the number of words be treated as figures or digits hereinafter.

On receiving the digit command from the business application, the control unit 5 enables the SRT 3 and PBR 4 (step 204) so as to recognize input signals from the channel control 1 at the same time, thereby automatically discriminating the input signals (step 205). When the predetermined number of figures have been recognized, the control unit 5 delivers the results of recognition to the business application 2 (step 206).

The automatic discrimination of input signals which is the characteristic feature of the present invention will be described in detail. Preconditions for the automatic identification are as follows:

(1) The detection rate of the PBR 4 is substantially 100 % while the recognition rate of the SRU 3 is less than 100 %; and

(2) There is no user who uses speech and PB together.

FIG. 3 shows the details of the automatic discrimination step 205 of FIG. 2 which the control unit 5 of the apparatus 10 executes. Implemented by a microprocessor, for example, the control unit 5 enables the SRU 3 and PBR 4 for the first digit in order to effect simultaneous recognition (step 204, FIG. 2). Then, the control unit 5 sets a predetermined time (T1) in a timer built therein. If the result of recognition of the first digit

is returned from the PBR 4 first (step 302), the control unit 5 immediately determines that it is the result of simultaneous recognition of the first digit (step 311) since the recognition rate of the PBR 4 is considered to be 100 %. At this instant, the control unit 5 disables the SRU 3. In the event of multi-figure input, the control unit 5 determines that the second and successive digits are not a speech on the basis of the previously stated precondition (2) and, therefore, executes the processing only with the PBR 4, i. e., without simultaneous recognition (step 312). When the result of recognition of the first digit is returned from the SRU 3 first (step 303), the control unit 5 waits a predetermined period of time (T2) to see if a result from the PBR 4 is not really returned. For this purpose, the control unit 5 sets the time T2 in a timer independent of the timer assigned to the time T1.

The time T2 should not be longer than about 1.5 seconds at most since it delays the processing time. If the PBR 4 returns an answer to the control unit 5 within the time T2 (step 308), the control unit 5 determines that the SRU 3 has misrecognized due to noise or similar cause. Then, the control unit 5 regards the result from the PBR 4 as the result of simultaneous recognition of the first digit and disables the SRU 3 (step 311). In the case of multi-figure input, the control unit 5 uses only the PBR 4 in effecting recognition (step 312). If an answer from the BPR 4 is not returned within the time T2 as determined

in the step 308, the control unit 5 determines that the input signal is a speech (step 309) and recognizes the second and successive digits only by the SRU 3 (step 310). On completing the recognition of the predetermined number of figures, the control unit 5 sends the results to the business application 2 (step 206). If the SRU 3 does not return an answer as determined in the step 303, the control unit 5 determines whether the time T1 has expired or not (step 304) and, if it has expired, ends the processing while informing the business application 2 of the expiration (step 306). If the time T1 has not expired as determined in the step 304, the program returns to the step 302 to see if the PBR 4 or the SRU 3 returns an answer.

By the above procedure, input signals are automatically discriminated.

As stated above, the illustrative embodiment recognizes the first digit and, base on the result of this recognition, recognizes the second and successive digits by either one of the SRU 3 or the PBR 4. This is successful so long as only the recognition unit associated with the input signal responds correctly. In practice, however, it sometimes occurs that both of the recognition units respond. Then, this embodiment effecting simultaneous recognition would malfunction. Generally, the two different recognition units may respond at the same time under either one of the following two situations:

- (a) The PBR 4 also responds to a speech input; and
- (b) The SRU 3 also responds to a PB input.

The above occurrence (a) is unavoidable although rare. Hence, considering that the probability that the occurrence (a) continues is low, it is determined that the input signal is PB if the PBR 4 responds within a predetermined plurality of digits. In the event of the occurrence (b), the input signal is determined to be PB since the recognition rate of the PBR 4 is considered to be 100 %. However, when PB involving noise or speech is recognized by the SRU 3, the PBR 4 is apt to return a result after the SRU 3. To eliminate this problem, a result from the PBR 4 may be waited for after the return of a result from the SRU 3. However, this implementation is not fully satisfactory since the waiting time delays the response and, therefore, cannot exceed a certain limit. A procedure which promotes more accurate automatic identification consists in determining, every time the SRU 3 returns a result, whether or not the PBR 4 returns an answer and regarding the input as PB if the PBR 3 has returned an answer as to two or more digits. FIG. 4 shows a sequence of steps for practicing such a procedure.

The procedure shown in FIG. 4 corresponds to the step 205 shown in FIG. 2. Specifically, both the SRU 3 and the PBR 4 are enabled. First, the control unit 5 sets in a first timer an input time T0 associated with the number of figures which is instructed by the business application 2. If the PBR 4 returns a

result of recognition of the first digit before the SRU 3 (step 402), the control unit 5 regards it as a result of simultaneous recognition on the first digit by considering that the recognition rate of the PBR 4 is 100 %. If it is the SRU 3 that has returned
5 a result first (step 403), the control unit 5 waits a predetermined period of time (T1) to see if the PBR 4 does not really return an answer (step 406). Again, this waiting time T1 should not be longer than about 1.5 second so as not to delay the processing. If the PBR 4 returns an answer within the period
10 of time T1, the control unit 5 regards the result from the PBR as a result of simultaneous recognition of the first digit by determining that the SRU 3 has misrecognized due to noise or similar cause (step 411).

If the first digit is PB as determined in the step 411, the
15 answer of a step 412 is NO without exception since the number of times that the PBR 4 answers is unconditionally once. Then, the next digit is recognized (step 413). At this time, the control unit 5 does not enable the SRU 3 and waits for a result from the PBR 4 for a given period of time (T2) (step 414). Assuming
20 that the user of the telephone 6 presses the keys on the telephone 6 slowly, the period of time T2 is the interval between the successive operations of the keys, e. g. 1 second to 3 seconds. If the PBR 4 returns a result within the period of time T2, the control unit 5 regards it as a result of simultaneous recognition
25 of the second digit (step 417). If the result of recognition from

the PBR is not on the last digit (step 418), the control unit 5 recognizes the succeeding digit or digits with the PBR 4 only, i. e., it does not execute simultaneous recognition (step 419). If the PBR 4 does not return a result within the period of time T2
5 as determined in the step 414 and if the SRU 3 has returned a result on the first digit (step 415), the control unit 5 replaces the result on the first digit with the result from the SRU (step 416) and enables the SRU 4 (step 410).

On the other hand, if the SRU 3 has returned a result on
10 the first digit first (step 403) and if the PBR 4 has not returned a result within the waiting time T1 (step 406), the control unit determines that the input is not PB since the PBR 4 is free from misrecognition. Then, the control unit 5 regards the result from the SRU 3 as a result of simultaneous recognition of the first
15 digit (step 407).

In the illustrative embodiment, when the answers from the SRU 3 and PBR 4 exist together, the results of recognition are replaced with each other, depending on the situation, as follows:

20 (i) First replacement: When the PBR 4 returns an answer as to two or more digits during the recognition of a plurality of digits, the results having been returned from the PBR 4 are substituted for the results of recognition; and

(ii) Second replacement: Assume that after a result
25 from the PBR 4 on a given digit has been regarded as a result of

2045959

recognition, the PBR 4 does not return an answer as to the next digit within the period of time T2. Then, if all the results of recognition up to the digit of interest are the results from the SRU 3, the results from the SRU 3 are substituted for the results
5 of recognition.

A reference will be made to FIGS. 5A through 5G for describing the answers from the SRU 3 and PBR 4 and the results of recognition on the assumption that five digits are sequentially recognized. As shown, when the SRU 3 answers first as to the
10 first digit (step 403) and the PBR 4 does not answer within the period of time T1 (step 406), the result S1 from the SRU 3 is determined to be the result of recognition of the first digit (step 407 and FIG. 5A). As the SRU 3 answers first as to the second digit also and the PBR 4 does not answer, the output S2 of the
15 SRU 3 is determined to be the result of recognition of the second digit (FIG. 5B). However, regarding the third digit, the PBR 4 answers before the SRU 3 (step 402), so that the output P1 of the PBR 4 is determined to be the result of recognition of the third digit (step 41 and FIG. 5C). Since the PBR 4 answers only
20 once, the step 412 is followed by the step 413 for recognizing the next digit. Assume that the PBR 4 has not returned an answer within the period of time T2 (step 414). Then, the control unit
5 determines whether or not the SRU 3 has responded to the third digit (step 415). If the answer of the step 415 is YES, the
25 above-mentioned replacement (i) is effected to substitute all of

B

2045959

the outputs of the SRU 3 having been returned from the results of recognition up to third digits (step 416 and FIG. 5D). Since the SRU 3 has to be enabled digit by digit, the control unit 5 enables it to process the next digit (step 410). Assume that the
5 PBR 4 has responded as to the next digit also (step 402). Then, the result from the PBR 4 is selected as a result of recognition of the digit of interest (step 411). Since the PBR has responded twice as counted from the beginning of the processing, the control unit 5 determines that the input signals
10 are PB and, therefore, effects the replacement (i) (steps 412 and 417 and FIG. 5E). After such a decision, the control unit 5 executes recognition with the succeeding digits by using only the PBR 4, i. e., without enabling the SRU 3 (step 419 and FIG. 5F). On determining the results of recognition of five digits
15 (step 418 and FIG. 5G), the control unit 5 sends them to the business application (step 206). Such a procedure enhances more accurate identification of input signals. It is to be noted that the number of digits to which the PBR responds as determined in the step 412 is not limited to two and may be three
20 or more.

In summary, it will be seen that the present invention provides a speech recognition apparatus which has a speech recognition unit for identifying a speech from an input signal, a PB recognition unit for detecting a PB signal from an input
25 signal, and a control unit capable of automatically determining

B

whether an input signal is a speech signal or whether it is a PB
signal. The apparatus, therefore, makes it needless for a
business application which controls it to discriminate a PB signal
and a speech signal. As a result, the customer intending to use
5 an inquiry service, for example, does not have to discriminate
the telephone number for voice input and the telephone number
for PB input. In the case of information service, it is not
necessary for the customer to register at the system regarding
the PB/voice input. Hence, the apparatus enhances the
10 serviceability of the system. Further, the system sets up
efficient traffic since both of PB input processing and voice input
processing are implemented by a single telephone channel,
exhibiting the processing ability to the full extend.

Various modifications will become possible for those
15 skilled in the art after receiving the teachings of the present
disclosure without departing from the scope thereof.

131

2045959

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE
PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

- 1 1. A speech recognition apparatus comprising:
 speech recognizing means for recognizing a speech from
 an input signal and outputting the result of recognition;
 PB signal detecting means for detecting a PB signal from
5 the input signal and outputting the result of detection; and
 control means for controlling said speech recognizing
 means and said PB signal detecting means to automatically
 determine whether the input signal is a speech signal or whether
 said input signal is a PB signal on the basis of the result of
10 recognition and the result of detection which said speech
 recognizing means and said PB signal detecting means output
 when used at the same time.
- 1 2. An apparatus as claimed in claim 1, wherein said
 control means determines that the input signal is a speech signal
 when a first response which is the response of said speech
 recognizing means precedes a second response which is the
5 response of said PB signal detecting means and if said second
 response does not appear within a predetermined period of time
 after said first response, or determines that said input signal is
 a PB signal when said second response appears within said
 predetermined period of time or when said second response
10 precedes said first response.
- 1 3. An apparatus as claimed in claim 1, wherein said
 control means determines that the input signal is a PB signal

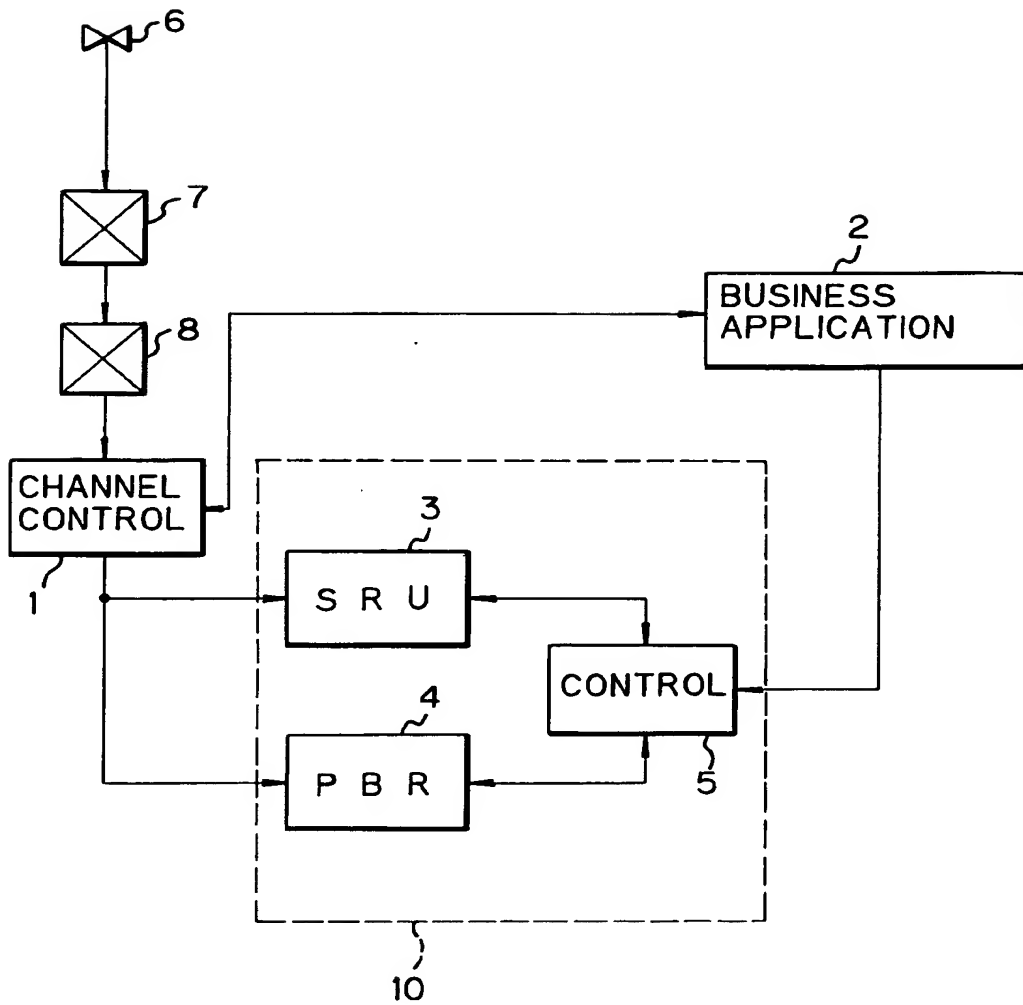
when a first response which is the response of said speech
recognizing means precedes a second response which is the
5 response of said PB signal detecting means and another second
response appears within a first predetermined period of time
after said second response, when said first response precedes
said second response and said second response appears within a
second period of time after said first response and another
10 second response appears within said first predetermined period
of time after said second response, or when said second
response appears more than a predetermined number of times
during a plurality of times of recognition of the input signal.



2045959

1/6

Fig. 1

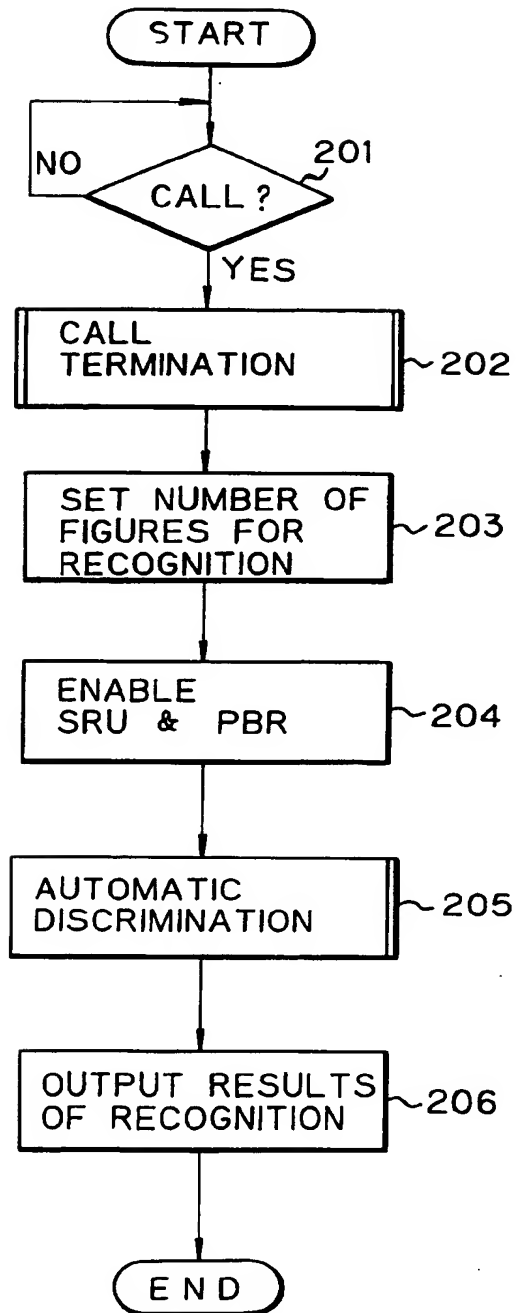


*Patent Agents
Smart & Biggar*

2045959

Fig. 2

2/6

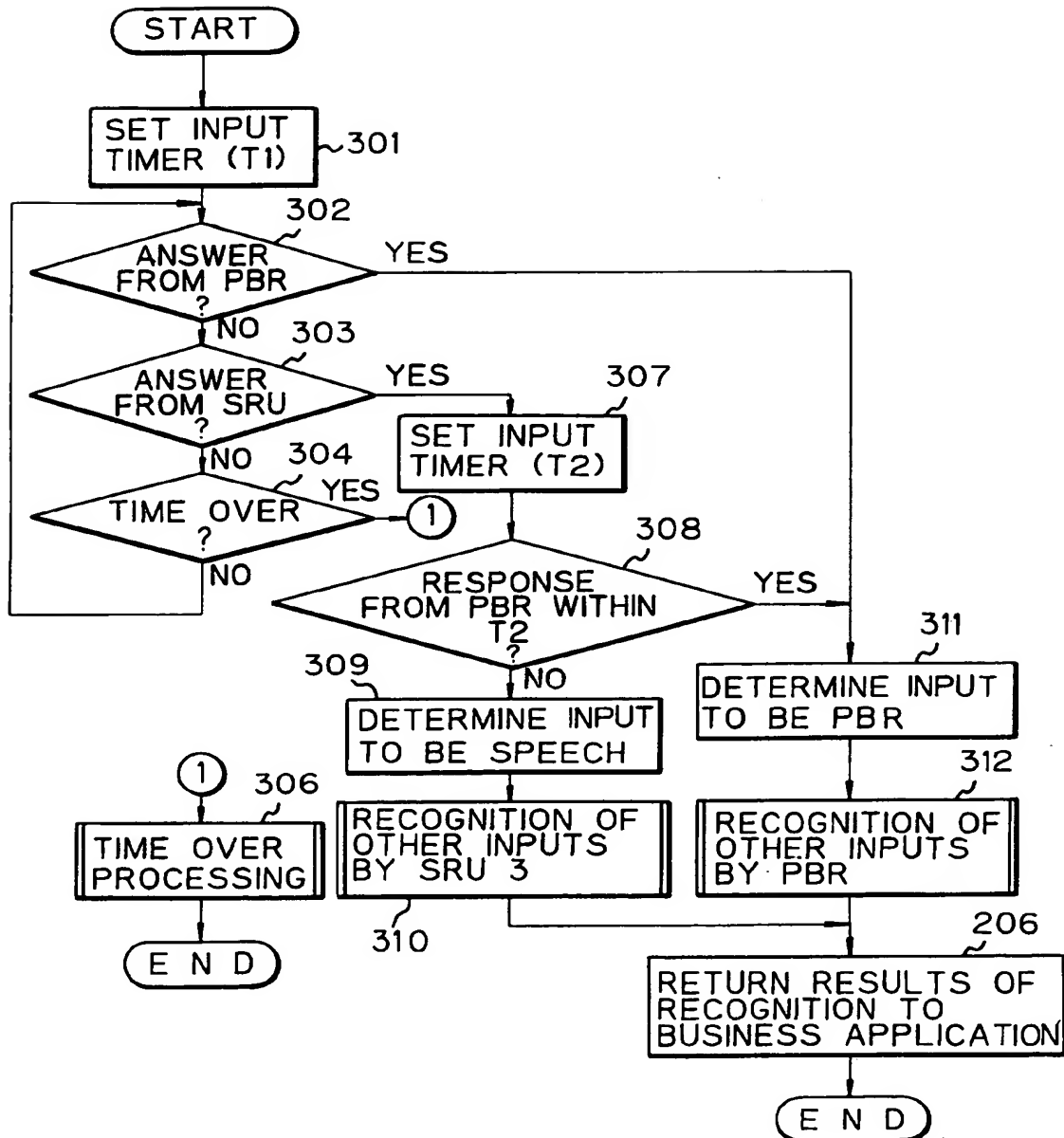


*Patent Agents
Smart & Biggar*

2045959

Fig. 3

3/6



Patent Agents
Smart & Biggar

Fig. 4
Fig. 4A
Fig. 4B

Fig. 4A

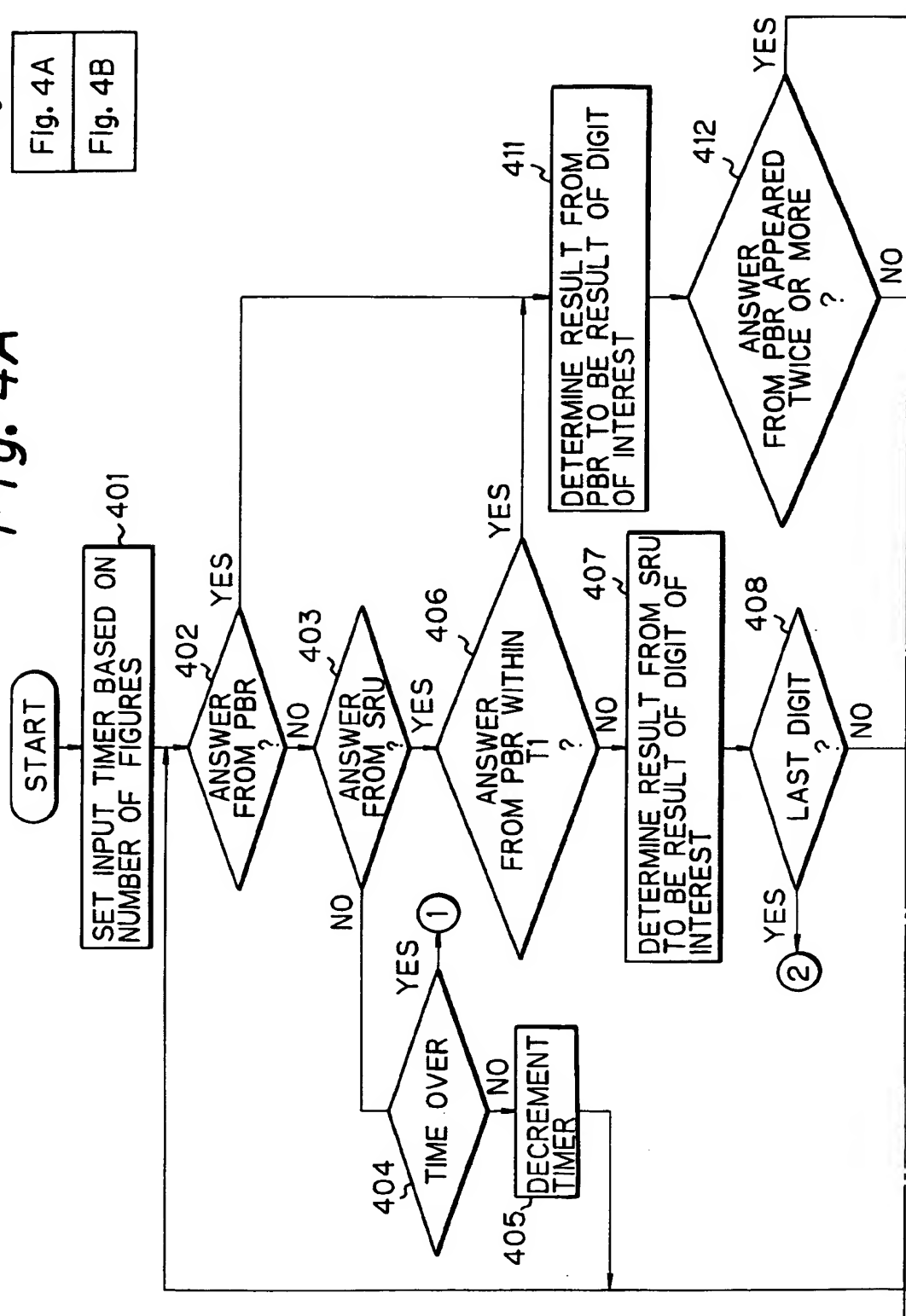


Fig. 4B

